International Journal of Networked and Distributed Computing, Vol. 3, No. 4 (November 2015), 214-223

An Energy Efficient Clustering Algorithm for Maximizing the Lifetime of Wireless Sensor Network

Kyung Tae Kim
College of Information & Communication Engineering, Sungkyunkwan University,
Suwon, 440-746*, Korea*
E-mail: kyungtaekim76@gmail.com

Man Youn Kim
College of Information & Communication Engineering, Sungkyunkwan University,
Suwon, 440-746*, Korea*
E-mail: {benimaru82}@skku.edu

Ji Hyeon Choi
College of Information & Communication Engineering, Sungkyunkwan University,
Suwon, 440-746*, Korea*
E-mail: {jichoi0530}@skku.edu

Hee Yong Youn
College of Information & Communication Engineering, Sungkyunkwan University,
Suwon, 440-746*, Korea*
E-mail: youn7147}@skku.edu

Abstract

In this paper we propose an energy efficient clustering algorithm for maximizing the lifetime of WSNs. The proposed scheme decides the cluster head based on the energy level of the nodes, while tree topology is adopted to connect the nodes inside each cluster. A new model for deciding an optimal number of clusters is also derived. Computer simulation reveals that the proposed scheme significantly extends the network lifetime and message delivery ratio compared to the existing schemes.

Keywords: Wireless sensor network; Energy efficiency; Network lifetime; Tree construction; Clustering.

1. Introduction

Advanced integrated circuit technologies have led to the development of small sensor nodes equipped with sensing, data processing and communication capability. Distributed in the target area, they form a network which can sense the environmental data and react to the surrounding condition. This makes the WSN suitable for a wide range of civil and military applications - target field imaging, intrusion detection, weather monitoring, security and tactical surveillance and disaster management, etc [1,2]. WSN is composed of a large number of sensor nodes and a base station (BS). The BS processes and stores the information it receives from the sensor nodes [3-5]. The sensor nodes are usually deployed randomly in the region of interest. A sensor node, despite its limited energy, processing capability and memory, collaborates with other sensor nodes, queries the physical environment, compiles the received data and transmits them to the BS. With their capabilities for monitoring and control, the network can provide a fine global picture of the target area through the integration of the data collected from many sensors each providing a coarse local view [6,7].

Since sensor nodes have limited power supply which cannot be recharged or replaced, their operation needs to be energy efficient. The limited energy in each node affects the lifetime of the entire network, and thus energy efficiency has been a critical design issue for the protocols and algorithms developed for WSN [8-10].

Cluster-based routing protocol is effective for prolonging the lifetime of WSN [11]. In cluster-based routing, the nodes in the network take different roles according to a variety of conditions and metric. Each cluster has a leader referred to as cluster-head (CH) and other ordinary member nodes. The CHs can form another hierarchy among them. The clustering approach allows a WSN of high scalability, less consumed energy and thus longer lifetime for the whole network [12]. This is mainly due to the fact that most of the sensing, data processing and communication activities can be performed within the clusters. However, energy consumption at a CH is significantly larger than that at other ordinary sensor nodes because CH is responsible for delivering aggregated data in its cluster to the BS. This problem can be relieved by rotating the role of CH among all nodes.

In this paper we propose an energy efficient clustering protocol employing tree topology for self-organizing WSN. The proposed scheme utilizes a new probability function to decide the number of CHs in each round. The function involves energy level, round information, frequency selected as CH of each node. The main focus of the proposed protocol is the minimization of energy consumption of the nodes for maximizing the network lifetime. For this, the proposed protocol adopts the tree topology inside each cluster. After leveling of the member nodes is completed with the CH as the root, a tree in a cluster is constructed by broadcasting the topology via a hybrid medium access control approach. The proposed scheme significantly reduces energy consumption and increase the lifetime of the sensor network compared to the existing schemes. Also, through the uniformly distributed CHs, the proposed scheme balances the energy consumption among the sensor nodes. The simulation results demonstrate that the proposed scheme effectively extends the network lifetime compared with the existing schemes.

The remainder of the paper is organized as follows. Section 2 presents the related work, and Section 3 introduces the proposed scheme. The performance of the proposed scheme is evaluated by computer simulation in Section 4. Finally, Section 5 concludes the paper and outlines future research direction.

2. Related Work

2.1. *Cluster-Based Routing in WSN*

The traditional routing protocols for WSN take energy efficiency as a primary criterion. Cluster-based routing has been recognized to be effective in this regards. It can improve the network scalability and lifetime, and also energy efficiency. It is efficient to lower the energy consumption within each cluster, performing data aggregation and fusion to decrease the number of transmitted messages to the BS [13-16].

Cluster-based routing is manipulated in rounds, where each round consists of four stages: cluster-head (CH) selection, cluster formation, intra-cluster communication and inter-cluster communication. In the first stage, the CHs are selected with the purpose of gathering data from the member nodes of its cluster. This stage is followed by the construction of the clusters. In other words, each node selects a CH to join. In the intra-cluster communication stage, the member nodes gather the data and send them to their CH. After receiving the data from the member nodes, the CHs perform data aggregation to omit redundant data. During the inter-cluster communication stage, the CHs send the data to the BS [17,18]. Both the intra and inter-cluster communication might be performed in single hop or multi-hop communication.

The clustering protocols have numerous merits which make them the most suitable and preferred protocols for WSN [19,20]. First, the total energy expenditure taken for data transmission is minimized. In addition, packet collision is reduced and redundant data are eliminated by the data aggregation process. Furthermore, the communication overhead to the BS is reduced since only the CHs participate in the data transmission. The bandwidth demand is lessened, while the limited bandwidth is used effectively. The overhead of routing and topology maintenance is small. Moreover, the network manageability and scalability are enhanced [21]. In next subsection various energy-efficient cluster-based routing protocols for WSN are discussed.

2.2. *Existing Schemes*

Low-Energy Adaptive clustering Hierarchy (LEACH) [22] is one of the representative clustering schemes. In LEACH the sensor nodes are divided into clusters, and a sensor node is selected as the CH in each cluster. The selection of CH is based on a predetermined probability.

Each node chooses a cluster to join by comparing the strengths of the advertisement messages sent from the CHs. The member nodes monitor the environment and send data to their CH. The CH is responsible for collecting and aggregating the information of the member nodes in the cluster before sending them to the BS. As the member nodes do not send data directly to the BS, the distance of data transmission is short, requiring small energy consumption.

The operation of LEACH consists of rounds. Each round begins with set-up phase in which the clusters are formed, followed by steady-state phase in which data are transferred from the nodes to the BS. During the set-up phase of cluster formation, each node decides whether it becomes a CH for the current round or not. The CHs are stochastically selected. Each node determines a random number between 0 and 1. If the number is smaller than a threshold, the node becomes a CH for the current round. The threshold for node-*n*, $T(n)$, is obtained by Eq. (1) below.

$$T(n) = \begin{cases} \dfrac{k}{N - k \times (r \bmod \dfrac{N}{k})} & , \forall n \in G \\ 0 & \forall n \notin G \end{cases} \quad (1)$$

where N is the total number of nodes, k is the number of clusters, r is the number of the current round, and G is the set of nodes that have not been CHs in the last N/k rounds. Using this approach, each node is ensured to be a CH exactly once within N/k rounds. After N/k rounds, all nodes are once again eligible to become CHs. As long as the energy efficiency is concerned, it is not desirable to select CHs randomly. However, repeating the rounds can offset the negative aspect of randomness.

After the CHs are selected, each CH broadcasts an advertisement message. The sensor nodes listen to the advertisements and join the closest one. Each CH then sets up a TDMA schedule for all the member nodes in its cluster. During the steady-state phase, the member nodes send the data they gather to the CH according to the TDMA schedule. The CH aggregates and compresses the data before passing them to the BS. At the end of a round, a new set of nodes become the CHs for the subsequent round and the whole process repeats.

In LEACH-C [30], the clusters are decided by the BS. In the set-up phase, the BS receives the messages containing the information on the location and amount of remaining energy of each sensor node. The BS then calculates the average energy of the nodes, and the ones having larger energy than the average are picked as candidate CHs. From the candidates, the nodes of smaller distance to the neighboring nodes are selected as CHs using the simulated annealing algorithm. After that, the BS broadcasts a message containing the IDs of the CHs for that round to the nodes in the network. When a node receives the message, it elects itself as CH if its identifier (ID) is included in the list of CHs. If not, the node determines its cluster and identifies the TDMA slot in which data are transmitted to the CH. The operation of LEACH-C in steady-state phase is identical to that of LEACH.

Proxy-Enable Adaptive Clustering Hierarchy for wireless sensor network (PEACH) [23] improved LEACH by selecting a proxy node which can assume the role of the current CH of weak power during one round of communication. In PEACH, the CH selection algorithm is the same as LEACH. Therefore, if all nodes are alive and have large energy, the number of CHs with PEACH will be the same as LEACH. However, in LEACH, CHs are randomly selected regardless of the energy remaining in the nodes. It may thus cause failing CHs and limit the lifetime of the sensor network. PEACH solves the problem of LEACH where the CH of no enough energy takes the duty of CH. If a cluster becomes to have a failing CH, a proxy is selected to operate in replace of the original CH. It is based on the consensus of healthy nodes for the detection and manipulation of failure of any CH. It allows improvement in the network lifetime by reducing the overhead of re-clustering.

The authors of [24] propose a protocol called Energy-Driven Adaptive Clustering Hierarchy (EDACH), which can increase the lifetime and reliability of sensor network in the presence of faults at the CHs. This is achieved by selecting a proxy node which can assume the role of the current CH during one round of communication. EDACH is based on consensus of healthy CHs to detect and handle faults in any faulty CH. EDACH employs the simulation-based fault injection method for performance evaluation, which assumes that errors occur according to a predetermined distribution. It provides improvement in the stability of the system and reduces the overhead of re-clustering and system reconfiguration.

[25] proposed a novel algorithm called EECH (Energy Efficient Clustering Hierarchy). Here the node of more energy has higher probability to be selected as

CH than other nodes. Besides, the CHs are set to use multi-hop forwarding and routing when they communicate with the BS. The EECHS scheme [26] adjusts the threshold value of the stochastic CH selection algorithm of LEACH. It considers residual energy of the nodes, distance between the nodes and the BS, and the number of consecutive rounds in which a node has not been a CH in CH selection. We next present the proposed scheme.

[27] proposed another enhancement over LEACH protocol. The protocol, called Power-Efficient Gathering in Sensor Information Systems (PEGASIS), is a near optimal chain-based protocol. The basic idea of the protocol is that the nodes communicate with only their closest neighbors and take turns to communicate with the BS to maximize the network lifetime. When the round of all nodes communicating with the BS ends, a new round begins. This reduces the power required to transmit data per round as the power draining is spread uniformly over all nodes. PEGASIS has two main objectives. First, increase the lifetime of each node with the collaboration between them, and as a result the network lifetime too. Second, allow collaboration only between the nodes close each other so that the bandwidth consumed for the communication is reduced. Unlike LEACH, PEGASIS avoids cluster formation and uses only one node in a chain to transmit to the BS instead of using multiple nodes. To locate the closest neighbor node, each node uses the signal strength to measure the distance to all neighboring nodes and then adjusts the signal strength so that only one node can be heard. The chain in PEGASIS consists of those nodes that are closest to each other, and forms a path to the BS. The aggregated data are sent to the BS by a node in the chain, and the nodes in the chain take turns in sending data to the BS. The chain construction is performed in a greedy fashion. PEGASIS increases the lifetime of network compared to the LEACH protocol. The performance gain is achieved through the elimination of the overhead caused by dynamic cluster formation in LEACH, and decreasing the number of transmissions and reception using data aggregation.

3. The Proposed Scheme

We first discuss the system model and energy model used in the proposed routing scheme.

3.1. *System Model and Energy Model*

We consider a WSN of a number of sensor nodes distributed randomly in the target area. The sensing nodes periodically form the clusters and have enough transmission power to reach the BS. The following assumptions on the sensor nodes and underlying network employed:

- All nodes are homogeneous and have the same capabilities. Each node is assigned a unique identifier (ID).
- All sensor nodes are started with the same initial energy.
- When two communicating sensor nodes are not within each other's radio range, data are forwarded through other nodes.
- A routing and MAC infrastructure are in place, and the communication environment is contention and error-free.
- Data fusion or aggregation is used to reduce the number of messages in the network.

We use the radio model adopted in [10-14], which is the first order radio model. To derive the energy consumption per any type of node, both the free space (d^2 power loss) and the multipath fading (d^4 power loss) channel model are used which are based on the distance between the transmitter and receiver. Power control is used to invert the loss by suitably configuring the power amplifier. For the communication between a non-cluster-head and its CH, the free space (*fs*) model is used. Between the CHs and the BS, the multipath (*mp*) model is used. The radios have power control and can expend minimal energy to reach the intended recipients. The radio can be turned off to avoid receiving unintended transmissions. The energy consumption model is described as follows. When a node transmits l bit over distance d, the energy it consumes is:

$$E_{Tx}(l,d) = E_{Tx\text{-}elec}(l) + E_{Tx\text{-}amp}(l,d) = \begin{cases} lE_{elec} + l\varepsilon_{fs}d^2, & d < d_0 \\ lE_{elec} + l\varepsilon_{mp}d^4, & d \geq d_0 \end{cases} \tag{2}$$

and a node receives l-bit data, energy it consumes is:

$$E_{Rx}(l) = E_{Rx\text{-}elec}(l) = lE_{elec} \tag{3}$$

Here E_{elec} is the unit energy consumed by the electronics to process one bit data, ε_{fs} and ε_{mp} are the amplifier factor for free-space and multi-path models, respectively, and d_0 is the reference distance to determine which model to use. In these models, a radio dissipates E_{elec}=50nJ/bit to run the transmitter or receiver circuitry, ε_{fs}=10$pJ/bit/m^2$ and ε_{mp}=0.0013$pJ/bit/m^4$ for the transmitter amplifier. For simplicity of calculation, we assume that the transmission range of each node is same on one condition that the transmission range should cover all the neighbors in the network. Also, we assume that all data packets contain the same number of bits and the energy for data aggregation is set as E_{DA} = 5$nJ/bit/signal$.

3.2. *The Proposed Approach*

In this subsection we introduce the proposed scheme which employs tree topology in each cluster to evenly distribute the energy load among the sensors in the network. The proposed scheme consists of two phases, clustering phase and data transmission phase. In the following sub-sections we discuss each of them in detail.

3.2.1. *Optimal Number of Clusters*

The cluster-based protocols adopting a simple hierarchical path selection approach do not need any information on the location of the nodes or upper layer control. Since each node is selected as CH with the same probability, the load can be balanced. However, the following issues need to be resolved.

Firstly, the optimum number of clusters, k_{opt}, needs to be decided. If the number of clusters is smaller than k_{opt}, some nodes may exhaust its energy for transmitting data to the CH locating far. With excessive number of clusters, on the other hand, the nodes will quickly deplete their energy for direct communication to the BS. LEACH sets k_{opt} as 5% of the nodes without any formal model.

Secondly, each node has equal probability to be a CH. If a node of low energy is selected as a CH, however, it will quickly deplete its energy due to the heavy load of CH. This shortens the network lifetime. Therefore, we need to introduce a new threshold value, which is decided based on the probability of optimal number of CHs and the residual energy of the nodes, to properly select the CHs.

During one round of operation, the energy consumption of a CH, E_{CH}, is due to three factors: data reception, data aggregation, and transmission of the aggregated data to the BS. Since the distance between the CH and BS is usually long, the multi-path (d^4 power loss) model is used, and E_{CH} becomes:

$$E_{CH} = lE_{elec}N_1 + lE_{DA}(N_1+1) + lE_{elec} + l_{mp}d^4_{toBS} \quad (4)$$

where l is the number of bits in each data packet. N_1 is the number of member nodes in a cluster having Poisson distribution, E_{DA} is the energy cost of data aggregation, d_{toBS} is the distance from the CH to the BS. With full data aggregation, each CH needs to process n/k_{opt} signals of length-l. The average number of member nodes in each cluster is [28]:

$$E\left[N_1 \mid N=n\right] \approx E\left[N_1\right] = \lambda_0/\lambda_1 \quad (5)$$

Here λ_0 and λ_1 denote the density of CHs and member nodes, respectively. $\lambda_1 = p\lambda$, $p = k/n$, and λ (=$\lambda_0 + \lambda_1$) is the density of Poisson distribution process. n(=$\lambda \times A$) is the number of nodes, while A is the size of the target area of square shape where the sensor nodes are deployed, k is the number of clusters, p is the probability for a node to be CH.

Without loss of generality, assume that the BS locates at the center of the target area. Hence the average distance from each CH to the BS for a square area of unit side length is:

$$\begin{aligned} E\left[D_1 \mid N=n\right] &= \iint D_1 \cdot P_A d_x d_y \\ &= \int_{-a}^{a}\int_{-a}^{a}\sqrt{x^2+y^2}\,\frac{1}{4a^2}d_x d_y = 0.765 \end{aligned} \quad (6)$$

Here D_1 is the variable of Poisson distribution denoting the distance from the CH whose coordinate is (x,y) to the BS. P_A is the probability density of CH in the area A.

According to Eq. (5) and (6), Eq. (4) can be expressed by:

$$E_{CH} = \frac{n-k}{k}lE_{elec} + \frac{n}{k}lE_{DA} + lE_{elec} + 0.342a^4 l\varepsilon_{mp} \quad (7)$$

Because each member node needs to transmit only l-bit data to its CH and the distance between them, d_{toCH}, is relatively short, the free space model is used. Thus the energy used in each member node is:

$$E_{non\text{-}CH} = lE_{elec} + l\varepsilon_{fs}d^2_{toCH} \quad (8)$$

d_{toCH} is [28]:

$$E\left[L_1 \mid N = n\right] \approx E\left[L_1\right] = \lambda_0 / 2\lambda_1^{3/2} \quad (9)$$

Here L_1 is the variable of Poisson distribution signifying the sum of the distance from the member node to the CH. Then, the average distance from the member node to its CH, $E[H_1]$, is:

$$E\left[H_1 \mid N = n\right] = \frac{E\left[L_1 \mid N = n\right]}{E\left[H_1 \mid N = n\right]} = 0.5\left(\frac{k\lambda}{n}\right)^{-\frac{1}{2}} \quad (10)$$

Therefore,

$$E_{non\text{-}CH} = lE_{elec} + \frac{n}{4k\lambda} l\varepsilon_{fs} \quad (11)$$

The energy dissipated in a cluster during a single round, $E_{cluster}$, and the total energy consumption, E_{total}, are:

$$E_{cluster} = lE_{CH} + N_1 E_{non\text{-}CH} \quad (12)$$

$$E_{total} = kE_{cluster} \quad (13)$$

According to Eq. (7), (11), (12), Eq. (13) can be expressed by:

$$E_{total} = l\left[\left(2n-k\right)\right]E_{elec} + nE_{DA} + 0.342d^4 k\varepsilon_{mp} + \frac{n\left(n-k\right)}{4k\lambda}\varepsilon_{fs} \quad (14)$$

Setting the derivative of E_{total} with respect to k to zero,

$$\frac{d}{dk}E_{total} = -E_{elec} + 0.342a^4\varepsilon_{mp} - \frac{n^2}{4k^2\lambda}\varepsilon_{fs} = 0 \quad (15)$$

Because $n = \lambda(4a^2)$, $\lambda = n/4a^2$. The optimum number of cluster, k_{opt}, and optimal probability for any node to be CH, p_{opt}, are then

$$k_{opt} = \sqrt{\frac{na^2\varepsilon_{fs}}{0.342a^4\varepsilon_{mp} - E_{elec}}} \quad (16)$$

$$p_{opt} = \frac{k_{opt}}{n} = \sqrt{\frac{a^2\varepsilon_{fs}}{n\left(0.342a^4\varepsilon_{mp} - E_{elec}\right)}} \quad (17)$$

3.2.2. *Selection of Cluster-Head*

LEACH adopts the random mechanism of CH selection, in which CH is randomly selected and all the sensor nodes within a cluster take turns to be the CH. This leads to balanced energy consumption of all nodes, and hence a longer lifetime of the network. However, this approach can locate the CHs unevenly. It causes the energy consumption of the nodes unbalanced and reduces the lifetime of the network. In order to solve this problem, the proposed scheme introduces three parameters to stochastically select the CH. The first parameter is the energy ratio between the current energy vs initial energy. The second one is the round number. The third one is the count that the node has been selected as the CH. The probability function of Node-n in the proposed scheme is expressed as follows with the three factors:

$$T(n) = \left(\frac{p_{opt}}{1 - p_{opt} \times (r \bmod \frac{1}{p_{opt}})}\left(\frac{E_{residual}}{E_{init}}\right)\right)\left(\frac{p_{opt}\sqrt{r}}{C_{ch} \bmod \frac{1}{p_{opt}}} + 1\right) \quad (18)$$

where p_{opt} is the probability to be CH, C_{ch} is the number of times the node has been selected as the CH until the current round, $E_{residual}$ is the residual energy, E_{init} is the initial energy, and r is the round number, respectively.

After the CHs are selected using the proposed probability function, each node electing itself as a CH for the current round broadcasts an advertisement message (*ADV_Msg*) to other nodes. When a node receives an *ADV_Msg* messages from the CHs, it sends the join-request message (*Join_REQ*) to the CH which it chooses as its CH based on the received signal power. After the CH receives the message (*Join_REQ*), the CHs identify their member nodes based on the received *Join_REQ* message. Once the clusters are created, tree configuration in each cluster begins.

3.2.3. *Tree Configuration in Cluster*

After the clusters are formed, a tree is built with the member nodes where the CH is the root. In the tree structure one sensor node can have one parent and many children nodes. All member nodes in a cluster are arranged in m levels starting from a CH. The CH is the root of tree, and it is at level 0.

In this phase each node selects a parent node. For this, the member nodes compute their tree level according to the distance from the CH. The distance between tree levels in a cluster, called communication radius (*CR*), is decided a priori based on node density in the target area considering the connection among the nodes. In [16], the critical transmission range for multi-hop connectivity was presented. The authors assume that the nodes are uniformly distributed in the field and each cell of size $C \times C$ contains at least one node. In this case, the network is guaranteed to be connected if the transmission range is $R_t = (1+\sqrt{5})C$. A cell in this context is defined as an area in the 2-dimensional space in which every node can communicate with every other node residing in every neighboring cell. In a clustered network, a cell can be defined as an area where every node can reach every other node residing in the same cell. According to [29], we obtain that CR is equal or longer than $(1+\sqrt{5})(\sqrt{L \times L/N})$, where L is the length of one side of a square field.

Each member node computes its tree level, L, as follows.

$$L = \left\lceil \frac{D_{CH}}{CR} \right\rceil + 1 \tag{19}$$

where D_{CH} is the distance to its CH based on the power of a signal received from the CH. Thus, D_{CH} can be obtained using the *RSSI* model. We also calculate the maximum tree level, L_{max}, as follows.

$$L_{max} = \frac{\text{The longest length of a field}}{CR} + 1 \tag{20}$$

After leveling of the member nodes is completed, tree inside a cluster is constructed by assigning a level to each member node. For the tree construction, this phase requires a number of iterations (steps), which we refer to as time interval T_i (i=1, 2, …, L_{max}). Here each step takes T_R, which should be long enough for the nodes of same tree level to communicate with the neighbor nodes. The path setup based on the level starts from the nodes of the smallest tree level, and terminates with the nodes of the level of L_{max}.

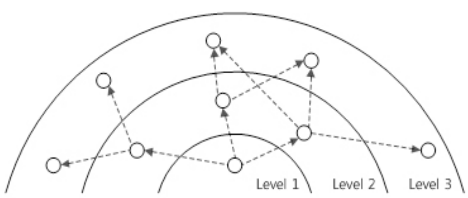


Fig. 1. Broadcast for tree construction in a cluster.

During iteration i, the nodes of level_i broadcast a message using CSMA, which is composed of the node's ID, level number, and ID of the parent node. Fig. 1 shows the broadcast process for tree construction in a cluster. Here each node decides a node as its parent node that transmits signal of the strongest power. When a node receives the reply from a node one level outer layer, it lists it as a child node. Then they turn off their radio until the data collection begins. The nodes of level_1 elect the CH as their parent node. The tree construction process continues until the tree covers all the nodes in the cluster.

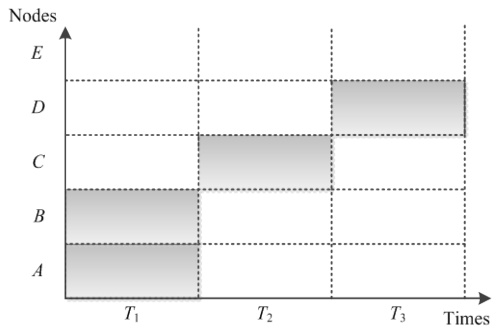


Fig. 2. The TDMA schedule in a cluster.

3.2.4. *Data Collection and Transmission*

Once the level-based tree construction in each cluster is complete, the data collection and transmission phase begins. In this phase each node sends the collected data to the parent node during the pre-allocated time slot appointed by the CH. We choose to implement a time division multiple access (TDMA)-based MAC layer for the slot assignment.

Each CH generates a TDMA schedule for the transmission and circulates it with the tree topology. In the allotted time-slot the nodes transmit gathered data to the lower level nodes. The lower level nodes fuse the received data with sensed data, and then send the result to the next lower level nodes. When the data from the nodes of all the levels in the cluster have been received, the CH applies data fusion to the received data. After this, it sends the fused data to the BS. Fig. 2 shows an example of data transmission schedule employed by the proposed scheme. In Fig. 2, node_A and B are located in the same level and node_C is lower level than them. The level of node_D is lower than that of node_C. Thus, in T_1, only node_ A and B transmit data to the parent nodes of the next level. The transmission process continues until the data transmission of all the nodes is completed.

4. Performance Evaluation

In this section we evaluate the performance of the proposed scheme using a simulator developed in C++, and compare the performance with the existing protocols. The simulation evaluates the number of CHs per round, the round number of *FND* (*first node die*) and *HNA* (*half of the nodes alive*), number of messages received, and the network lifetime with the protocols compared. Table 1 provides the parameters used in the simulation.

Table 1. The parameters used in the simulation.

Parameter	Value
Network size	$100m \times 100m$, $200m \times 200m$
Location of BS	Center of target area
Number of nodes	100
Data packet size	500 byte
Packet size for path set-up	20 byte
Network topology	Random
Initial energy of each sensor	1.0J
E_{elec}	50 nJ/bit
ε_{fs}	10 pJ/bit/ m^2
ε_{mp}	0.0013 pJ/bit/ m^4
E_{DA}	5 nJ/bit/signal

In the simulation the BS is assumed to be fixed and located at the center of the target area, and it has unlimited resources including the power. In order to study how the proposed scheme works with the network of different densities, the simulations are conducted with two different sizes of network of 100m × 100m and 200m × 200m. 1000 simulation runs are executed, and the results are averaged. In the simulations the radio model presented in Section 3.1 is adopted. The free space propagation model is used when the propagation distance is smaller than the threshold distance d_0. Otherwise, the ground reflection (two-ray) propagation model is used [22,24,26]. For the simplicity, an ideal MAC layer and error-free communication links are assumed.

To investigate the effectiveness of cluster formation of the proposed approach, we first examine the performance of CH selection. Fig. 3 shows the distribution of the number of clusters in LEACH, EECHS, and the proposed scheme, which is achieved from 100 rounds of the simulation. In Fig. 3, it is apparent that the number of clusters with the proposed scheme is more stable than that with LEACH and EECHS. LEACH uses a fully random approach to select the CHs with a pre-determined probability. As a result, the number of CHs with LEACH and EECHS varies significantly as shown in Fig. 3. The selection of CH with the proposed scheme is made considering several operation parameters of each node. Therefore, the proposed scheme allows more effective and consistent formation of clusters than LEACH and EECHS.

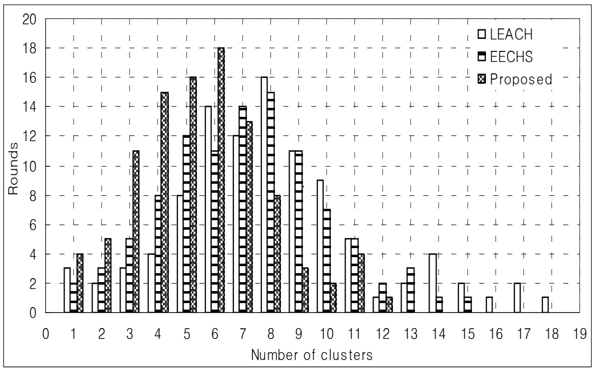


Fig. 3. The distribution of the numbers of clusters in each round.

We compare the proposed scheme with LEACH, EDACH, and EECHS in terms of network lifetime using FND and HNA. Fig. 4 shows that the proposed scheme improves FND approximately 70%, 51%, and 17% than LEACH, EDACH and EECHS, respectively.

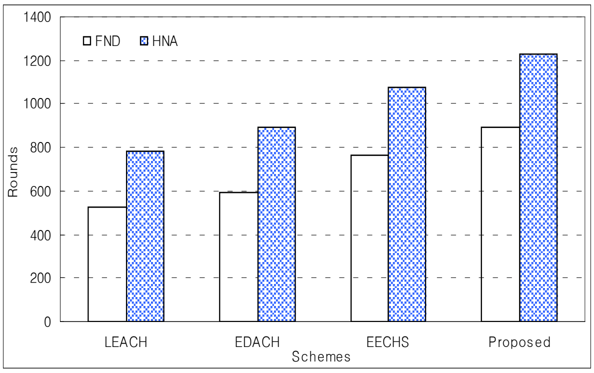
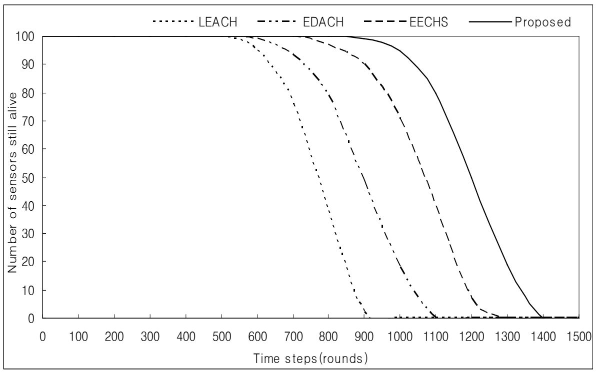


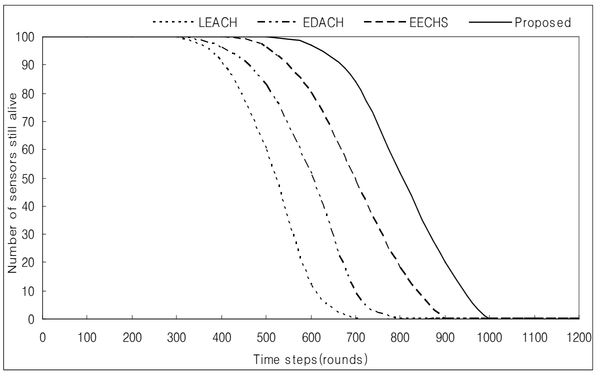
Fig. 4. The comparison of the schemes in terms of FND and HNA.

The improvements of HNA are approximately 58%, 40%, and 16%, respectively. Based on these results, the proposed scheme can be said to be better than the existing schemes in balancing the energy consumption of the sensor nodes.

We also evaluate the energy efficiency of the schemes by examining the network lifetime. Fig. 5 shows the number of sensor nodes still alive over the simulation time. The proposed scheme clearly outperforms the others with respect to the network lifetime. In Fig. 5(a), 100 nodes are randomly placed in a 100*m* × 100*m* area with the BS located at (50, 50). Fig. 5(b) is the simulation result with 200*m* × 200*m* area. Note that the existing schemes select the CHs based on a random approach. The proposed scheme allows considerable improvement in the network lifetime by conspicuous selection of the CHs and tree formation inside the clusters. Also, the proposed scheme outperforms other protocols even for larger size network.



(a) 100*m*×100*m*



(b) 200*m*×200*m*

Fig. 5. The comparison of network lifetimes.

Fig. 6 shows the total number of messages received by the BS over time with different schemes. For this experiment, we consider a network of 100 sensor nodes randomly distributed in a place of 100*m* × 100*m* network, where each node begins with an initial energy of 1.0J. The figure clearly illustrates the effectiveness of the proposed scheme in delivering significantly more messages than its counterparts. This was achieved by deciding optimal number of CHs and forming a tree in each cluster.

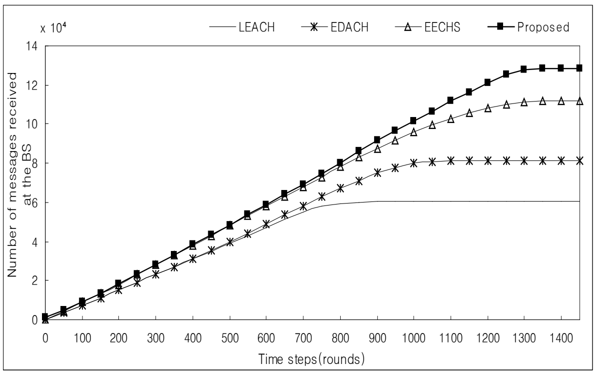


Fig. 6. The Total number of message received by the BS.

5. Conclusion

In this paper we have proposed an energy efficient clustering protocol for maximizing the network lifetime of wireless sensor network. The proposed scheme introduces a new threshold value used in selecting the CHs in the network. We have also introduced a new tree construction approach inside each cluster to minimize the energy consumption of the sensor nodes. As a result, the proposed scheme can significantly reduce energy consumption and increase the lifetime of the network compared to the existing schemes. Through the uniformly distributed CHs, the proposed scheme can effectively balance the energy consumption among all sensor nodes.

In the future we will continue to study the maximization of network lifetime for heterogeneous sensor network. In addition, mobility in sensor network is an ever-growing requirement in recent applications. Extension of the proposed scheme to cope with the mobility and the related challenges is yet another important issue remaining as future work.

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012R1A1A2040257 and 2013R1A1A2060398), the second Brain Korea 21 PLUS project, ICT R&D program of MSIP/IITP

(1391105003), and Samsung Electronics (S-2014-0700-000). Corresponding author: Hee Yong Youn.

References

1. I. F. Akyildiz et al., "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, March 2002.
2. K. T. Kim and H. Y. Youn, "A Stochastic and Optimized Energy Efficient Clustering Protocol for Wireless Sensor Networks," *International Journal of Distributed Sensor Networks*, vol.2014, pp.1–12, March 2014.
3. A. E. Tumer and M. Gunduz, "An improved leach protocol for indoor wireless sensor networks," in *Proc.* SPIN, pp. 432–437, Feb. 2014,
4. D. Culler, D. Estrin, and M. Srivastava, "Overview of sensor networks," *IEEE Computer*, pp. 41–49, 2004.
5. J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer Networks*, vol. 52, no. 12, pp. 2292–2330, Aug. 2008.
6. R. A. Gupta and M.-Y. Chow, "Networked control system: Overview and research trends," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 7, pp. 2527–2535, 2010.
7. J. A. Khna, H. K. Qureshi, and A. lqbal, "Energy management in Wireless Sensor Networks: A survey," *Computer & Electrical Engineering*, pp. 159–176, 2015.
8. Y. A. Obaisat and R. Braun, "On Wireless Sensor Networks: Architectures, Protocols, Applications, and Management," *Institute of Information and Communication Technologies*, pp.1–11, 2007.
9. P. Kumar, M. P. Singh, and U. S. Triar, "A Review of Routing Protocols in Wireless Sensor Network," *International Journal of Engineering Research & Technology*, vol. 1, no. 4, pp. 1–14, June 2012.
10. C. Li, H. X. Zhang, B. B. Hao, and J. D. Li, "A survey on routing protocols for large-scale wireless sensor networks," *Sensors*, vol. 11, pp. 3498-3526, 2011.
11. H. Deng, C. Yang, and Y. Sun, "A Novel Algorithm for Optimized Cluster Head Selection," *Science Journal of Electrical & Electronic Engineering*, Vol.2013, pp. 1–9, 2013.
12. S. K. Singh, M. P. Singh, and D. K. Singh, "Routing Protocols in Wireless Sensor Networks – A survey," *International Journal of Computer Science & Engineering Survey*, vol. 1, pp.60–64, 2010.
13. P. Kumar, M. P. Singh, and U. S. Triar, "A Review of Routing Protocols in Wireless Sensor Network," *International Journal of Engineering Research & Technology*, vol. 1, pp. 1–14, 2012.
14. S. Saranya and M. Princy, "Routing Techniques in Sensor Network–A Survey," *Procedia Engineering*, vol. 38, pp. 2739–2747, 2012.
15. M. Maimour, H. Zeghilet and F. Lepage, "Cluster-based Routing Protocol for Energy Efficiency in Wireless Sensor Networks," *Sustainable Wireless Sensor Networks*, pp. 167–188, 2010.
16. N. Rathi, J. Saraswa and P. P Bhattacharya, "A Review on Routing Protocols for Application in Wireless Sensor Network," *International Journal of Distributed and Parallel System*, vol. 3, no. 5, pp. 39–58, 2012.
17. X. Liu, "A Survey on Clustering Routing Protocols in Wireless Sensor Network," *Sensors*, vol. 12, pp. 11113–11153, 2012.
18. N. Nokhanji and Z. M. Hanapi, "A survey on cluster-based routing protocols in wireless sensor networks," *Journal of Applied Sciences*, vol. 14, pp. 2011–2022, 2014.
19. M. Haneef and Z. Deng, "Design challenges and comparative analysis of cluster based routing protocols used in wireless sensor networks for improving network lifetime," *Adv. Inf. Sci. Serv. Sci.*, vol. 4, pp. 450–459, 2012.
20. Naeimi, H. Ghafghazi, C. O. Chow, and H. Ishii, "A survey on the taxonomy of cluster-based routing protocols for homogeneous wireless sensor networks," *Sensor*, vol. 12, no. 6, pp. 7350–7409, 2012.
21. X. Liu and J. Shi, "Clustering routing algorithms in wireless sensor networks: An overview," *KSII Transactions on Internet and Information Systems*, vol. 6, pp. 1735–1755, 2012.
22. W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Micro-sensor Networks," In *Proc. HICSS*, pp.1–10, Jan. 2000.
23. K. T. Kim and H. Y. Youn, "PEACH: Proxy-Enable Adaptive Clustering Hierarchy for Wireless Sensor network," in *Proc. ICWN*, pp. 52–57, June 2005.
24. K. T. Kim and H. Y. Youn, "Energy-Driven Adaptive Clustering Hierarchy (EDACH) for Wireless Sensor Networks," *LNCS*, vol. 3823, pp. 1098–1107, Dec. 2005.
25. Y. Hu, W. Li, Z. Kang, "Study on Energy Efficient Hierarchical Routing Protocols of Wireless Sensor Network," in *Proc. ICIE*, pp.325–328, 2009.
26. A. Ray and D. De "Energy Efficient Clustering Hierarchy Protocol for Wireless Sensor Network," in *Proc. ICCIA*, pp.1–4, Dec. 2011.
27. S. Lindsey, C. S. Raghavendra, "PEGASIS: Power-Efficient gathering in sensor information systems," in *Proceedings of the IEEE Aerospace Conf.*, Canada, March 2002, pp. 1–6.
28. S. G. Foss and S. A. Zuyev, "On a Voronoi Aggregative Process Related to a Bivariate Poisson Process," *Advances in Applied Probability*, pp. 965-981, 1996.
29. O. Younis and S. Fahmy, "HEED: A Hybrid, Energy-Efficient, Distributed clustering approach for Ad Hoc sensor networks," *IEEE Trans. Mobile Computing*, vol. 3, no. 4, pp. 366–379, 2004.
30. W. B. Heinzelman, A. Chandrakasan, and H. Balakrishanan, "An Application-Specific Protocol Architecture for Wireless Microsensor Networks," *IEEE Transactions on Wireless Communications*, vol. 1, pp. 660–670, 2002.